(will be inserted by hand later)

Smooth maps from clumpy data: Covariance analysis

Marco Lombardi and Peter Schneider

Institüt für Astrophysik und Extraterrestrische Forschung, Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany

Received ***date***; Accepted ***date***

Abstract. Interpolation techniques play a central role in Astronomy, where one often needs to smooth irregularly sampled data into a smooth map. In a previous article (2001AA...373..359L, hereafter Paper I), we have considered a widely used smoothing technique and we have evaluated the expectation value of the smoothed map under a number of natural hypotheses. Here we proceed further on this analysis and consider the variance of the smoothed map, represented by a two-point correlation function. We show that two main sources of noise contribute to the total error budget and we show several interesting properties of these two noise terms. The expressions obtained are also specialized to the limiting cases of low and high densities of measurements. A number of examples are used to show in practice some of the results obtained.

Key words. methods: statistical – methods: analytical – methods: data analysis – gravitational lensing

1. Introduction

Raw astronomical data are very often discrete, in the sense that measurements are performed along a finite number of directions on the sky. In many cases, the discrete data are believed to be single measurements of a smooth underlying field. In such cases, it is desirable to reconstruct the original field using interpolation techniques. A typical example of the general situation just described is given by weak lensing mass reconstructions in clusters of galaxies. In this case, thousands of noisy estimates of the tidal field of the cluster (shear) can be obtained from the observed shapes of background galaxies whose images are distorted by the gravitational field of the cluster. All these measurements can then be combined to produce a smooth map of the cluster shear, which in turn is subsequently converted into a projected density map of the cluster mass distribution.

One of the most widely used interpolation techniques in Astronomy is based on a weighted average. More precisely, a positive weight function, describing the relative weight of a datum at the position $\boldsymbol{\theta} + \boldsymbol{\phi}$ on the point $\boldsymbol{\theta}$, is introduced. The weight function is often chosen to be of the form $w\bigl(|\boldsymbol{\phi}|\bigr)$, i.e. depends only on the separation $|\boldsymbol{\phi}|$ of the two points considered. Normally, w is also a decreasing function of $|\boldsymbol{\phi}|$ in order to ensure that the largest contributions to the interpolated value at $\boldsymbol{\theta}$ comes from nearby measurements. Then, the data are averaged using a weighted mean with the weights given by the function w. More precisely, calling $\hat{f}_n$ the n-th datum obtained at the position $\boldsymbol{\theta}_n$, the smooth map is defined as

$$\tilde{f}(\boldsymbol{\theta}) \equiv \frac{\sum_{n=1}^{N} \hat{f}_n w(\boldsymbol{\theta} - \boldsymbol{\theta}_n)}{\sum_{n=1}^{N} w(\boldsymbol{\theta} - \boldsymbol{\theta}_n)} \; , \tag{1}$$

where N is the total number of objects. In a previous paper (2001AA...373..359L, hereafter Paper I) we have evaluated the expectation value for this expression under the following hypothesis:

– The measured values $\{\hat{f}_n\}$ are independent random variables with expectation value

$$\bigl\langle \hat{f}_n \bigr\rangle = f(\boldsymbol{\theta}_n) \; . \tag{2}$$

In other words, the $\{\hat{f}_n\}$ are *unbiased* measurements of a field $f(\boldsymbol{\theta})$.

– The positions $\{\boldsymbol{\theta}_n\}$ are independent random variables with uniform distribution and density ρ. In Paper I we initially considered a fixed number N of positions inside a field Ω of finite area A; then, we took the *continuous limit* letting N go to infinity with $\rho = N/A$ constant. Equivalently, we considered N to be a Poissonian random variable with average ρA:

$$p_N(N) = \mathrm{e}^{-\rho A} \frac{(\rho A)^N}{N!} \; , \tag{3}$$

and each location $\boldsymbol{\theta}_n$ to be uniformly distributed inside A:

$$p_\theta(\boldsymbol{\theta}_n) = \frac{1}{A} \; . \tag{4}$$

Send offprint requests to: M. Lombardi
Correspondence to: lombardi@astro.uni-bonn.de

In Paper I we have shown that

$$\bigl\langle \tilde{f}(\boldsymbol{\theta}) \bigr\rangle = \int f(\boldsymbol{\theta}') w_{\mathrm{eff}}(\boldsymbol{\theta} - \boldsymbol{\theta}') \, \mathrm{d}^2\theta' \; . \tag{5}$$

Thus, $\langle \tilde{f} \rangle$ is the convolution of the unknown field f with an *effective weight* w_{eff} which, in general, differs from the weight function w. We also have shown that w_{eff} has a "similar" shape as w and converges to w when the object density ρ is large; however for finite ρ, w_{eff} is broader than w.

Here we proceed further with the statistical analysis by obtaining an expression for the two-point correlation function (covariance) of this estimator. More precisely, given two points $\boldsymbol{\theta}_A$ and $\boldsymbol{\theta}_B$, we will consider the two-point correlation function of $\tilde{f}$, defined as

$$\mathrm{Cov}(\tilde{f}; \boldsymbol{\theta}_A, \boldsymbol{\theta}_B) \equiv \bigl\langle \tilde{f}(\boldsymbol{\theta}_A) \tilde{f}(\boldsymbol{\theta}_B) \bigr\rangle - \bigl\langle \tilde{f}(\boldsymbol{\theta}_A) \bigr\rangle \bigl\langle \tilde{f}(\boldsymbol{\theta}_B) \bigr\rangle \tag{6}$$

In our calculations, similarly to Paper I, we will assume that $\hat{f}_n$ are *unbiased and mutually independent* estimates of $f(\boldsymbol{\theta}_n)$ [cf. Eq. (2)]. We will also assume that the $\{\hat{f}_n\}$ have fixed variance σ^2, so that

$$\bigl\langle \bigl[\hat{f}_n - f(\boldsymbol{\theta}_n) \bigr] \bigl[\hat{f}_m - f(\boldsymbol{\theta}_m) \bigr] \bigr\rangle = \sigma^2 \delta_{nm} \; . \tag{7}$$

The paper is organized as follows. In Sect. 2 we derive the general expression for the covariance of the interpolating techniques and we show that two main noise terms contribute to the total error. These results are then generalized in Sect. 3 to include the case of weight functions that are not strictly positive. A useful expansion at high densities ρ of the covariance is obtained in Sect. 4. Section 5 is devoted to the study of several interesting properties of the expressions obtained in the paper. Finally, in Sect. 6 we consider three simple weight functions and derive (analytically or numerically) the covariance for these cases. Three appendixes on more technical topics complete the paper.

2. Evaluation of the covariance

2.1. Preliminaries

Before starting the analysis, let us introduce a simpler notation. In the following we will often drop the arguments $\boldsymbol{\theta}_A$ and $\boldsymbol{\theta}_B$ in $\mathrm{Cov}(\tilde{f}; \boldsymbol{\theta}_A, \boldsymbol{\theta}_B)$ and other related quantities. Note, in fact, that the problem is completely defined with the introduction of the two "shifted" weight functions $w_A(\boldsymbol{\theta}) \equiv w(\boldsymbol{\theta}_A - \boldsymbol{\theta})$ and $w_B(\boldsymbol{\theta}) \equiv w(\boldsymbol{\theta}_B - \boldsymbol{\theta})$. We also call $\tilde{f}_A \equiv \tilde{f}(\boldsymbol{\theta}_A)$ and $\tilde{f}_B \equiv \tilde{f}(\boldsymbol{\theta}_B)$ the values of $\tilde{f}$ at the two points of interest $\boldsymbol{\theta}_A$ and $\boldsymbol{\theta}_B$, so that

$$\tilde{f}_X = \frac{\sum_{n=1}^N \hat{f}_n w_X(\boldsymbol{\theta}_n)}{\sum_{n=1}^N w_X(\boldsymbol{\theta}_n)} \; . \tag{8}$$

Hence, Eq. (6) can be rewritten in this notation as

$$\mathrm{Cov}(\tilde{f}) = \bigl\langle \tilde{f}_A \tilde{f}_B \bigr\rangle - \bigl\langle \tilde{f}_A \bigr\rangle \bigl\langle \tilde{f}_B \bigr\rangle \; . \tag{9}$$

Note that, using this notation, we are not taking advantage of the invariance upon translation of $w(\boldsymbol{\theta})$ in Eq. (1); in other words, we are not using the fact that w_A and w_B are basically the same function shifted by $\boldsymbol{\theta}_A - \boldsymbol{\theta}_B$. Actually, all calculations can be carried out without using this property; however, we will explicitly point out simplifications that can be made using the invariance upon translation.

We would also like to spend a few words about averages. Note that, as anticipated in Sect. 1, we need to carry out two averages, one with respect to $\{\hat{f}_n\}$ [Eqs. (2) and (7)], and one with respect to $\{\boldsymbol{\theta}_n\}$ [Eqs. (3) and (4)]. Taking $\{\boldsymbol{\theta}_n\}$ to be random variables is often reasonable because in Astronomy one has not a direct control over the positions where observations are made (think for example of weak lensing, where the data are represented by galaxy ellipticities); it has also the advantage of letting us to obtain general results, independent of any particular configuration of positions. Note, however, that taking $\{\boldsymbol{\theta}_n\}$ to be *independent* variables is a strong simplification which might produce inaccurate results in some context (**?**, see, e.g.,)]LPM. Finally, since the number of observations N is itself a random variable, we need to perform first the average on $\{\hat{f}_n\}$ and then the one on $\{\boldsymbol{\theta}_m\}$.

In closing this section, we observe that in this paper, similarly to Paper I, we will almost always consider the smoothing problem on the plane, i.e. we will assume that the positions $\{\boldsymbol{\theta}_n\}$ are vectors of $\mathbb{R}^2$. We proceed this way because in Astronomy the smoothing process often takes places on small regions of the celestial sphere, and thus on sets that can be well approximated with subsets of the plane. However, we stress that all the results stated here can be easily applied to smoothing processes that takes places on different sets, such as the real axis $\mathbb{R}$ or the space $\mathbb{R}^3$.

2.2. Analytical solution

Let us now focus on the first term on the r.h.s. of Eq. (9). We have

$$\bigl\langle \tilde{f}_A \tilde{f}_B \bigr\rangle = \frac{1}{A^N} \int_\Omega \mathrm{d}^2\theta_1 \int_\Omega \mathrm{d}^2\theta_2 \cdots \times \int_\Omega \mathrm{d}^2\theta_N \frac{\bigl\langle \bigl[\sum_n \hat{f}_n w_A(\boldsymbol{\theta}_n) \bigr] \bigl[\sum_m \hat{f}_m w_B(\boldsymbol{\theta}_m) \bigr] \bigr\rangle}{\bigl[\sum_n w_A(\boldsymbol{\theta}_n) \bigr] \bigl[\sum_m w_B(\boldsymbol{\theta}_m) \bigr]} \; . \tag{10}$$

Note that the average in the r.h.s. of this equation is only with respect to $\{\hat{f}_n\}$. Expanding the numerator in the integrand of this equation, we obtain N^2 terms, N of which have $n = m$ and $N(N-1)$ have $n \neq m$. We can then rewrite Eq. (10) above as

$$\bigl\langle \tilde{f}_A \tilde{f}_B \bigr\rangle = T_1 + T_2 \; , \tag{11}$$

where

$$T_1 \equiv \frac{1}{A^N} \int_\Omega \mathrm{d}^2\theta_1 \int_\Omega \mathrm{d}^2\theta_2 \cdots \times \int_\Omega \mathrm{d}^2\theta_N \frac{\sum_n \langle \hat{f}_n^2 \rangle w_A(\boldsymbol{\theta}_n) w_B(\boldsymbol{\theta}_n)}{\bigl[\sum_n w_A(\boldsymbol{\theta}_n)\bigr] \bigl[\sum_m w_B(\boldsymbol{\theta}_m)\bigr]} , \tag{12}$$

$$T_2 \equiv \frac{1}{A^N} \int_\Omega \mathrm{d}^2\theta_1 \int_\Omega \mathrm{d}^2\theta_2 \cdots \times \int_\Omega \mathrm{d}^2\theta_N \frac{\sum_{n \neq m} \langle \hat{f}_n \hat{f}_m \rangle w_A(\boldsymbol{\theta}_n) w_B(\boldsymbol{\theta}_m)}{\bigl[\sum_n w_A(\boldsymbol{\theta}_n)\bigr] \bigl[\sum_m w_B(\boldsymbol{\theta}_m)\bigr]} . \tag{13}$$

Despite the apparent differences, these two terms can be simplified in a similar manner. Let us consider first T_1. Using Eq. (7), we can evaluate the average $\langle \hat{f}_n^2 \rangle = \sigma^2 + \bigl[f(\boldsymbol{\theta}_n)\bigr]^2$. Since the positions $\{\boldsymbol{\theta}_n\}$ appear as "dummy variables" in Eq. (12), we can relabel them as follows

$$T_1 = \frac{N}{A^N} \int_\Omega \mathrm{d}^2\theta_1 \int_\Omega \mathrm{d}^2\theta_2 \cdots \times \int_\Omega \mathrm{d}^2\theta_N \frac{\bigl[f^2(\boldsymbol{\theta}_1) + \sigma^2\bigr] w_A(\boldsymbol{\theta}_1) w_B(\boldsymbol{\theta}_1)}{\bigl[\sum_n w_A(\boldsymbol{\theta}_n)\bigr] \bigl[\sum_m w_B(\boldsymbol{\theta}_m)\bigr]} . \tag{14}$$

In order to simplify this equation, we use a technique similar to the one adopted in Paper I. More precisely, we split the two sums in the denominator of the integrand of Eq. (14), taking away the terms $w_A(\boldsymbol{\theta}_1)$ and $w_B(\boldsymbol{\theta}_1)$. Hence, we write

$$T_1 = \frac{1}{\rho} \int_\Omega \mathrm{d}^2\theta_1 \bigl[f^2(\boldsymbol{\theta}_1) + \sigma^2\bigr] w_A(\boldsymbol{\theta}_1) w_B(\boldsymbol{\theta}_1) \times C\bigl(w_A(\boldsymbol{\theta}_1), w_B(\boldsymbol{\theta}_1)\bigr) , \tag{15}$$

where $C(w_A, w_B)$ is a *corrective factor* given by

$$C(w_A, w_B) \equiv \frac{N^2}{A^{N+1}} \int_\Omega \mathrm{d}^2\theta_2 \cdots \int_\Omega \mathrm{d}^2\theta_N \times \frac{1}{\bigl[w_A + \sum_{n=2}^N w_A(\boldsymbol{\theta}_n)\bigr] \bigl[w_B + \sum_{m=2}^N w_B(\boldsymbol{\theta}_n)\bigr]} . \tag{16}$$

The additional factor $\rho = N/A$ has been introduced to simplify some of the following equations. Note that in the definition of C w_A and w_B are formally taken to be two real variables (instead of two real functions of argument $\boldsymbol{\theta}_1$).

The definition of C above suggests to define two new random variables y_A and y_B:

$$y_X \equiv \sum_{n=2}^N w_X(\boldsymbol{\theta}_n) , \qquad \text{with } X = \{A, B\} . \tag{17}$$

Note that the sum runs from $n = 2$ to $n = N$. If we could evaluate the *combined* probability distribution function $p_y(y_A, y_B)$ for y_A and y_B, we would have solved our problem: In fact we could use this probability to write $C(w_A, w_B)$ as follows

$$C(w_A, w_B) = \rho^2 \int_0^\infty \mathrm{d}y_A \int_0^\infty \mathrm{d}y_B \frac{p_y(y_A, y_B)}{(w_A + y_A)(w_B + y_B)} . \tag{18}$$

To obtain the probability distribution $p_y(y_A, y_B)$, we need to use the combined probability distribution $p_w(w_A, w_B)$ for w_A and w_B. This distribution is implicitly defined by saying that the probability that $w_A(\boldsymbol{\theta})$ be in the range $[w_A, w_A + \mathrm{d}w_A]$ and $w_B(\boldsymbol{\theta})$ be in the range $[w_B, w_B + \mathrm{d}w_B]$ is $p_w(w_A, w_B) \, \mathrm{d}w_A \, \mathrm{d}w_B$. We can evaluate $p_w(w_A, w_B)$ using

$$p_w(w_A, w_B) = \frac{1}{A} \int_\Omega \delta\bigl(w_A - w_A(\boldsymbol{\theta})\bigr) \delta\bigl(w_B - w_B(\boldsymbol{\theta})\bigr) \, \mathrm{d}^2\theta . \tag{19}$$

Turning back to (y_A, y_B), we can write a similar expression for p_y:

$$p_y(y_A, y_B) = \frac{1}{A^{N-1}} \int_\Omega \mathrm{d}^2\theta_2 \cdots \int_\Omega \mathrm{d}^2\theta_N \delta(y_A - w_{A2} - \cdots - w_{AN}) \times \delta(y_B - w_{B2} - \cdots - w_{BN}) , \tag{20}$$

where for simplicity we have called $w_{Xn} = w_X(\boldsymbol{\theta}_n)$. Note that inserting this equation into Eq. (18) we recover Eq. (16), as expected. Actually, for our purposes it is more useful to consider y_X to be the sum of N random variables $\{w_{Xn}\}$. In other words, we consider the set of couples $\bigl\{(w_{An}, w_{Bn})\bigr\}$, made of the two weight functions at the various positions, as a set of N *independent* two-dimensional random variables (w_A, w_B) with probability distribution $p_w(w_A, w_B)$. [Hence, similarly to Eq. (16), we consider the weight functions w_X to be real variables instead of real functions; the independence of the positions $\boldsymbol{\theta}_n$ then implies the independence of the *couples* (w_{An}, w_{Bn}).] Taking this point of view, we can rewrite Eq. (20) as

$$p_y(y_A, y_B) = \int_0^\infty \mathrm{d}w_{A2} \int_0^\infty \mathrm{d}w_{B2} \, p_w(w_{A2}, w_{B2}) \cdots \times \int_0^\infty \mathrm{d}w_{AN} \int_0^\infty \mathrm{d}w_{BN} \, p_w(w_{AN}, w_{BN}) \times \delta(y_A - w_{A2} - \cdots - w_{AN}) \delta(y_B - \cdots - w_{BN}) . \tag{21}$$

It is well known in Statistics that the sum of independent random variables with the same probability distribution can be better studied using Markov's method (see, e.g., 1989QB461.C47......; see also 1987PhRvL..59.2814D for an application to microlensing studies). This method is based on the use of Fourier transforms for the probability distributions p_w and p_y. However, since we are dealing with non negative quantities (we recall that we assumed $w(\boldsymbol{\theta}) \geq 0$), we can replace the Fourier transform with Laplace transform which turns out to be more appropriate in for our problem (see Appendix C for a summary of the properties of Laplace transforms). Hence, we define $W(s_A, s_B)$ and $Y(s_A, s_B)$ to be the Laplace transforms of $p_w(w_A, w_B)$ and $p_y(w_A, w_B)$, respectively. Note that, since both functions p_w and p_y have two arguments, we need two arguments for the Laplace transforms as well:

$$W(s_A, s_B) \equiv \mathcal{L}[p_w](s_A, s_B) = \int_0^\infty \mathrm{d}w_A \int_0^\infty \mathrm{d}w_B \mathrm{e}^{-s_A w_A - s_B w_B} p_w(w_A, w_B) , \tag{22}$$

$$Y(s_A, s_B) \equiv \mathcal{L}[p_y](s_A, s_B) = \int_0^\infty \mathrm{d}y_A \int_0^\infty \mathrm{d}y_B \mathrm{e}^{-s_A y_A - s_B y_B} p_y(y_A, y_B) . \tag{23}$$

We use now in these expressions the Eq. (19) for p_w and Eq. (21) for p_y, thus obtaining

$$W(s_A, s_B) = \frac{1}{A} \int_\Omega \mathrm{e}^{-s_A w_A(\boldsymbol{\theta}) - s_B w_B(\boldsymbol{\theta})} \, \mathrm{d}^2\theta , \tag{24}$$

$$\begin{aligned} Y(s_A, s_B) &= \frac{1}{A^{N-1}} \int_\Omega \mathrm{d}^2\theta_2 \cdots \int_\Omega \mathrm{d}^2\theta_N \\ &\quad \times \exp\left[-s_A \sum_{n=2}^N w_{An} - s_B \sum_{m=2}^N w_{Bm} \right] \\ &= \bigl[W(s_A, s_B) \bigr]^{N-1} . \end{aligned} \tag{25}$$

Hence, p_y can in principle be obtained from the following scheme. First, we evaluate $W(s_A, s_B)$ using Eq. (24), then we calculate $Y(s_A, s_B)$ from Eq. (25), and finally we back-transform this function to obtain $p_y(y_A, y_B)$.

Actually, another, more convenient, technique is viable. Following the path of Paper I, we now take the "continuous limit" and treat N as a random variable. As explained in Sect. 1, we can take this limit using two equivalent approaches:

- We keep the area A fixed and consider N to be a random variable with Poisson distribution given by Eq. (3). We then average over all possible configurations obtained.
- We take the limit $N \to \infty$ taking the density $\rho = N/A$ fixed.

We will follow here the second strategy. In the limit $A \to \infty$ the quantity $W(s_A, s_B)$ goes to unity and thus is not useful for our purposes. Instead, it is convenient to define

$$Q(s_A, s_B) \equiv \int_\Omega \bigl[\mathrm{e}^{-s_A w_A(\boldsymbol{\theta}) - s_B w_B(\boldsymbol{\theta})} - 1 \bigr] \, \mathrm{d}^2\theta = A \bigl[W(s_A, s_B) - 1 \bigr] . \tag{26}$$

This definition is sensible because, this way, Q remains finite for $A \to \infty$. In the continuous limit, Eq. (25) becomes

$$Y(s_A, s_B) = \lim_{N \to \infty} \left[1 + \frac{Q(s_A, s_B)\rho}{N} \right]^{N-1} = \mathrm{e}^{\rho Q(s_A, s_B)} . \tag{27}$$

In order to evaluate $C(w_A, w_B)$, we rewrite its definition (18) as

$$C(w_A, w_B) = \rho^2 \int_0^\infty \mathrm{d}x_A \int_0^\infty \mathrm{d}x_B \frac{\zeta_w(x_A, x_B)}{x_A x_B} , \tag{28}$$

where $x_X \equiv y_X + w_X$ and

$$\zeta_w(x_A, x_B) \equiv \mathrm{H}(x_A - w_A) \mathrm{H}(x_B - w_B) p_y(x_A - w_A, x_B - w_B) . \tag{29}$$

Here $\mathrm{H}(x_X - w_X)$ are Heaviside functions at the positions w_X, i.e.

$$\mathrm{H}(x) = \begin{cases} 0 & \text{if } x < 0 , \\ 1 & \text{otherwise.} \end{cases} \tag{30}$$

Note that ζ_w is basically a "shifted" version of p_y. Looking back at Eq. (28), we can interpret the integration present in this equation as a *very* particular case of Laplace transform with vanishing argument. In other words, we can write

$$C(w_A, w_B) = \rho^2 \mathcal{L}[\zeta_w / x_A x_B](0, 0) . \tag{31}$$

Thus our problem is solved if we can obtain the Laplace transform of $\zeta_w / x_A x_B$ evaluated at $s_A = s_B = 0$. From the properties of Laplace transform [cf. Eq. (C.7)] we find

$$\mathcal{L}[\zeta_w(x_A, x_B)/x_A x_B](s_A, s_B) = \int_{s_A}^\infty \mathrm{d}s'_A \int_{s_B}^\infty \mathrm{d}s'_B \, Z_w(s'_A, s'_B) , \tag{32}$$

where Z_w is the Laplace transform of ζ_w:

$$Z_w(s_A, s_B) \equiv \mathcal{L}[\zeta_w](s_A, s_B) = \mathrm{e}^{-s_A w_A - s_B w_B} Y(s_A, s_B) . \tag{33}$$

Combining together Eqs. (31), (32), and (33) we finally obtain

$$C(w_A, w_B) = \rho^2 \int_0^\infty \mathrm{d}s_A \int_0^\infty \mathrm{d}s_B \, \mathrm{e}^{-s_A w_A - s_B w_B} Y(s_A, s_B) = \rho^2 \mathcal{L}[Y](w_A, w_B) . \tag{34}$$

In summary, the set of equations that can be used to evaluate T_1 are

$$Q(s_A, s_B) = \int_\Omega \bigl[\mathrm{e}^{-s_A w_A(\boldsymbol{\theta}) - s_B w_B(\boldsymbol{\theta})} - 1 \bigr] \, \mathrm{d}^2\theta . \tag{35}$$

$$Y(s_A, s_B) = \exp\bigl[\rho Q(s_A, s_B) \bigr] . \tag{36}$$

$$C(w_A, w_B) = \rho^2 \int_0^\infty \mathrm{d}s_A \int_0^\infty \mathrm{d}s_B \, \mathrm{e}^{-s_A w_A - s_B w_B} Y(s_A, s_B) = \rho^2 \mathcal{L}[Y](w_A, w_B) . \tag{37}$$

$$T_1 = \frac{1}{\rho} \int_\Omega \mathrm{d}^2\theta \bigl[f^2(\boldsymbol{\theta}) + \sigma^2 \bigr] w_A(\boldsymbol{\theta}) w_B(\boldsymbol{\theta}) \times C\bigl(w_A(\boldsymbol{\theta}), w_B(\boldsymbol{\theta}) \bigr) . \tag{38}$$

These equations solve completely the first part of our problem, the determination of T_1.

Let us now consider the second term of Eq. (11), namely T_2 [see Eq. (13)]. We first evaluate the average in $\{\hat{f}_n\}$ that appears in the numerator of the integrand of Eq. (13), obtaining $\langle \hat{f}_n \hat{f}_m \rangle = f(\boldsymbol{\theta}_n) f(\boldsymbol{\theta}_m)$ [cf. Eq. (7) with $n \neq m$]. Then we relabel the "dummy" variables

$\{\boldsymbol{\theta}_n\}$ similarly to what has been done for T_1, thus obtaining

$$T_2 = \frac{N(N-1)}{A^N} \int_\Omega \mathrm{d}^2\theta_1 \int_\Omega \mathrm{d}^2\theta_2 \cdots \times \int_\Omega \mathrm{d}^2\theta_N \frac{f(\boldsymbol{\theta}_1) w_A(\boldsymbol{\theta}_1) f(\boldsymbol{\theta}_2) w_B(\boldsymbol{\theta}_1)}{\bigl[\sum_n w_A(\boldsymbol{\theta}_n)\bigr] \bigl[\sum_m w_B(\boldsymbol{\theta}_m)\bigr]} . \tag{39}$$

We now split, in the two sums in the denominator, the terms $w_A(\boldsymbol{\theta}_1) + w_A(\boldsymbol{\theta}_2)$ and $w_B(\boldsymbol{\theta}_1) + w_B(\boldsymbol{\theta}_2)$ and define the new random variables

$$z_X \equiv \sum_{n=3}^N w_X(\boldsymbol{\theta}_n) , \qquad \text{with } X = \{A, B\} . \tag{40}$$

Again, if we know the *combined* probability distribution $p_z(z_A, z_B)$ of z_A and z_B our problem is solved, since we can write [cf. Eqs. (15) and (18)]

$$T_2 = \frac{N(N-1)}{A^2} \int_\Omega \mathrm{d}^2\theta_1 \int_\Omega \mathrm{d}^2\theta_2\, f(\boldsymbol{\theta}_1) f(\boldsymbol{\theta}_2) w_A(\boldsymbol{\theta}_1) w_B(\boldsymbol{\theta}_2) \times \int_0^\infty \mathrm{d}z_A \int_0^\infty \mathrm{d}z_B\, p_z(z_A, z_B) \frac{1}{w_A(\boldsymbol{\theta}_1) + w_A(\boldsymbol{\theta}_2) + z_A} \times \frac{1}{w_A(\boldsymbol{\theta}_1) + w_B(\boldsymbol{\theta}_2) + z_B} . \tag{41}$$

Actually, in the continuous limit, z_X is indistinguishable from y_X (z_X differs from y_X only on the fact that it is the sum of $N-2$ "weights" instead of $N-1$; however, N goes to infinity in the continuous limit and thus y_X and z_X converge to the same quantity). Thus we can rewrite Eq. (41) as

$$T_2 = \int_\Omega \mathrm{d}^2\theta_1 \int_\Omega \mathrm{d}^2\theta_2\, f(\boldsymbol{\theta}_1) f(\boldsymbol{\theta}_2) w_A(\boldsymbol{\theta}_1) w_B(\boldsymbol{\theta}_2) \times C\bigl(w_A(\boldsymbol{\theta}_1) + w_A(\boldsymbol{\theta}_2), w_B(\boldsymbol{\theta}_1) + w_B(\boldsymbol{\theta}_2)\bigr) , \tag{42}$$

where C is still given by Eq. (37).

Finally, in order to evaluate $\mathrm{Cov}(\tilde{f})$, we still need the simple averages $\langle \tilde{f}_A \rangle$ and $\langle \tilde{f}_B \rangle$. These can be obtained directly using the technique described in Paper I, where we have shown that the set of equations to be used is

$$Q_X(s_X) \equiv \int_\Omega \bigl[\mathrm{e}^{-s_X w_X(\boldsymbol{\theta})} - 1\bigr] \,, \mathrm{d}^2\theta , \tag{43}$$

$$Y_X(s_X) \equiv \exp\bigl[\rho Q_X(s_X)\bigr] , \tag{44}$$

$$C_X(w_X) \equiv \rho \int_0^\infty \mathrm{d}s_X \mathrm{e}^{-s_X w_X} Y_X(s_X) . \tag{45}$$

$$\langle \tilde{f}_X \rangle = \int_\Omega \mathrm{d}^2\theta\, f(\boldsymbol{\theta}) w_X(\boldsymbol{\theta}) C_X\bigl(w_X(\boldsymbol{\theta})\bigr) . \tag{46}$$

We recall that in Paper I we called the combination $w_{\mathrm{eff}X}(\boldsymbol{\theta}) = w_X(\boldsymbol{\theta}) C_X\bigl(w_X(\boldsymbol{\theta})\bigr)$ *effective weight*. Alternatively, the correcting factors $C_X(w_X)$ can also be obtained from $C(w_A, w_B)$ using the following equations

$$C_A(w_A) \equiv \lim_{w_B \to \infty} C\bigl(w_A, w_B\bigr) , \tag{47}$$

$$C_B(w_B) \equiv \lim_{w_A \to \infty} C\bigl(w_A, w_B\bigr) . \tag{48}$$

We now have at our disposal the complete set of equations that can be used to determine the covariance of $\tilde{f}$.

In closing this subsection we makes a few comments on the translation invariance for w_X (see Sect. 2.1). Since $w_A(\boldsymbol{\theta})$ and $w_B(\boldsymbol{\theta})$ are basically the same function with a shift, the two functions Q_A and Q_B are the same, so that C_A coincides with C_B. Not surprisingly, the two effective weights $w_{\mathrm{eff}A}$ and $w_{\mathrm{eff}B}$ differ also only by a shift.

2.3. Noise contributions

A simple preliminary analysis of the Eqs. (38) and (42) allows us to recognize two main sources of noise. In fact, a term in Eq. (38) is proportional to σ^2, and is clearly related to measurement errors of f, namely

$$T_\sigma \equiv \frac{\sigma^2}{\rho} \int_\Omega \mathrm{d}^2\theta w_A(\boldsymbol{\theta}) w_B(\boldsymbol{\theta}) C\bigl(w_A(\boldsymbol{\theta}), w_B(\boldsymbol{\theta})\bigr) . \tag{49}$$

Other factors entering $\mathrm{Cov}(\tilde{f})$ can be interpreted as Poisson noise. Hence, we call $T_{\mathrm{P1}} \equiv T_1 - T_\sigma$, $T_{\mathrm{P2}} \equiv T_2$, and $T_{\mathrm{P3}} \equiv \langle \tilde{f}_A \rangle \langle \tilde{f}_B \rangle$, so that the total Poisson noise is $T_{\mathrm{P}} \equiv T_{\mathrm{P1}} + T_{\mathrm{P2}} - T_{\mathrm{P3}}$.

The noise term T_σ is quite intuitive and does not require a long explanation. We note here only that this term is independent of the field $f(\boldsymbol{\theta})$ because we assumed measurements $\hat{f}_n$ with fixed variance σ^2 [see Eq. (7)].

The Poisson noise T_{P} can be better understood with a simple example. Suppose that $f(\boldsymbol{\theta})$ is *not* constant and let us focus on a point where this function has a strong gradient. Then, when measuring $\tilde{f}$ in this point, we could obtain an excess of signal because of an overdensity of objects in the region where $f(\boldsymbol{\theta})$ is large; the opposite happens if we have an overdensity of objects in the region where $f(\boldsymbol{\theta})$ is small. This noise source, called Poisson noise, vanishes if the function $f(\boldsymbol{\theta})$ is flat.

In the rest of this paper we will study the properties of the two-point correlation function. Before proceeding, however, we need to consider an important generalization of the results obtained here to the case of vanishing weights.

3. Vanishing weights

So far we have implicitly assumed that both w_A and w_B are always positive. In some cases, however, it might be interesting to consider vanishing weight functions (for example, functions with finite support). We need then to modify accordingly our equations.

When using vanishing weights, we might encounters situations where the denominator of Eq. (1) vanishes. In this case, clearly, the estimator $\tilde{f}$ cannot be even defined, and any further statistical analysis is meaningless. This problem was already encountered in Paper I, where we used the following prescription. If we are using a finite-field weight function, we mark as "bad configurations" for the point $\boldsymbol{\theta}$ the sets $\{\boldsymbol{\theta}_n\}$ for which $\tilde{f}(\boldsymbol{\theta})$ is not defined.

Then, in taking the ensemble average for $\tilde{f}(\boldsymbol{\theta})$, we explicitly exclude the bad configurations $\{\boldsymbol{\theta}_n\}$. The same prescription will be also adopted to evaluate the covariance of our estimator. Hence, when performing the ensemble average to estimate the covariance $\mathrm{Cov}(\tilde{f}; \boldsymbol{\theta}_A, \boldsymbol{\theta}_B)$, we explicitly exclude configurations where either $\tilde{f}_A$ or $\tilde{f}_B$ cannot be evaluated. This is implemented with a slight change in the definition of p_y, which in turn implies a change in Eq. (36) for Y. A rigorous generalization of the relevant equations can now be carried out without significant difficulties. However, the equations obtained are quite cumbersome and present some technical peculiarities. Hence, we prefer to postpone a complete discussion of vanishing weights until Appendix A; we report here only the main results.

As mentioned above, the basic problem of having vanishing weights is that in some cases the estimator is not defined. Hence, it is convenient to define three probabilities, namely P_A and P_B, the probabilities, respectively, that $\tilde{f}_A$ and $\tilde{f}_B$ are not defined, and P_{AB}, the probability that both quantities are not defined. Note that, because of the invariance upon translation for w, we have $P_A = P_B$. These probabilities can be estimated without difficulties. In fact, the quantity $\tilde{f}_X$ is not defined if and only if there is no object inside the support of w_X. Since the number of points inside the support of w_X follows a Poissonian probability, we have $P_X = \exp(-\rho \pi_X)$, where π_X is the area of the support of w_X. Similarly, calling $\pi_{A\cap B}$ the area of the intersection of the supports of w_A and w_B, we find $P_{AB} = \exp(-\rho \pi_{A\cap B})$. Using Eqs. (35) and (36) we can also verify the following relations

$$P_{AB} = \lim_{\substack{s_A \to \infty \\ s_B \to \infty}} Y(s_A, s_B) , \tag{50}$$

$$P_A = \lim_{\substack{s_A \to 0^+ \\ s_B \to \infty}} Y(s_A, s_B) , \tag{51}$$

$$P_B = \lim_{\substack{s_B \to \infty \\ s_B \to 0^+}} Y(s_A, s_B) . \tag{52}$$

$$1 = \lim_{\substack{s_B \to 0^+ \\ s_B \to 0^+}} Y(s_A, s_B) . \tag{53}$$

Appendix A clarifies better the relationship between the limiting values of Y and the probabilities defined above. In the following we will use a simplified notation for limits, and we will write something like $P_A = Y(0^+, \infty)$ for Eq. (51).

The only significant modification to the equations obtained above for vanishing weights is an overall factor in Eq. (37), which now becomes

$$C(w_A, w_B) = \frac{\rho^2}{1 - P_A - P_B + P_{AB}} \mathcal{L}[Y](w_A, w_B) . \tag{54}$$

The factor $1/(1 - P_A - P_B + P_{AB})$ is basically a renormalization; more precisely, it is introduced to take into account the fact that we are discarding cases where either $\tilde{f}_A$ or $\tilde{f}_B$ are not defined. Note, in fact, that in agreement with the inclusion-exclusion principle, $(1 - P_A - P_B + P_{AB})$ is the probability that the both $\tilde{f}_A$ and $\tilde{f}_B$ are defined. Since the combination $(1 - P_A - P_B + P_{AB})$ enters several equations, we define

$$\nu \equiv \frac{1}{1 - P_A - P_B + P_{AB}} . \tag{55}$$

Equation (54) is the most important correction to take into account for vanishing weights. Actually, there are also a number of peculiarities to consider when dealing with the probability p_y and its Laplace transform Y. Fortunately, however, these peculiarities have no significant consequence for our purpose and thus we can still safely use Eqs. (35) and (36). Again, we refer to Appendix A for a complete explanation.

4. Moments expansion

In most applications, the density of objects is rather large. Hence, it is interesting to obtain an expansion for $C(w_A, w_B)$ valid at high densities.

In Paper I we already obtained an expansion for $C_A(w_A)$ (or, equivalently, $C_B(w_B)$) for $\rho \to \infty$:

$$C_A(w_A) \simeq \frac{\rho}{\rho + w_A} + \frac{\rho^2 S_{20}}{(\rho + w_A)^3} - \frac{\rho^2 S_{30}}{(\rho + w_A)^4} + \frac{\rho^2 S_{40} + 3\rho^3 S_{20}^2}{(\rho + w_A)^5} . \tag{56}$$

In this equation, S_{ij} are the moments of the functions (w_A, w_B), defined as

$$S_{ij} \equiv \int_\Omega \mathrm{d}^2\theta \bigl[w_A(\boldsymbol{\theta})\bigr]^i \bigl[w_B(\boldsymbol{\theta})\bigr]^j . \tag{57}$$

Clearly, in Eq. (56) enter only the moments S_{i0}, since the form of w_B is not relevant for $C_A(w_A)$. Similarly, the expression for $C_B(w_B)$ contains only the moments S_{0j}. Note that for weight functions invariant upon translation we have $S_{ij} = S_{ji}$.

A similar expansion can be obtained for $C(w_A, w_B)$. Calculations are basically a generalization of what was done in Paper I for $C(w)$ and can be found in Appendix B. Here we report only the final result obtained:

$$C(w_A, w_B) \simeq \frac{\rho^2}{(\rho + w_A)(\rho + w_B)} + \frac{\rho^3 S_{20}}{(\rho + w_A)^3 (\rho + w_B)} + \frac{\rho^3 S_{11}}{(\rho + w_A)^2 (\rho + w_B)^2} + \frac{\rho^3 S_{02}}{(\rho + w_A)(\rho + w_B)^3} . \tag{58}$$

Figure 1 shows the results of applying this expansion to a Gaussian weight. For clarity, we have considered in this figure (and in others shown below) a 1-dimensional smoothing instead of the 2-dimensional case discussed in the text, and we have used x as spatial variable instead of $\boldsymbol{\theta}$. The figure refers to two identical Gaussian weight functions with vanishing average and unit variance. A comparison of this figure with Fig. 2 of Paper I shows that the convergence here is much slower. Nevertheless, Eq. (58) will be very useful to investigate some important limiting cases in the next section.

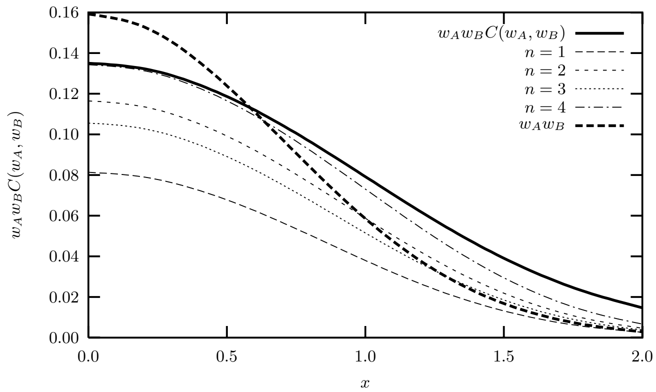

Fig. 1. The moment expansion of $C(w_A, w_B)$ for 1-dimensional Gaussian weight functions $w_A(x) = w_B(x)$ centered on 0 and with unit variance. The plot shows the various order approximations obtained using the method described in Sect. 4 (equations for the orders $n = 3$ and $n = 4$ are not explicitly reported in the text; see however Table B.1 in Appendix B). The density used is $\rho = 1$.

5. Properties

In this section we will study in detail the two noise terms T_σ and T_P introduced in Sect. 2.3, showing their properties in several limiting cases.

5.1. Normalization

A simple normalization property for $C(w_A, w_B)$ can be derived, similarly to what we have already done for the average of $\tilde{f}$ in Paper I. Suppose that $f(\boldsymbol{\theta}) = 1$ and that no errors are present on the measurements, so that $\sigma^2 = 0$. In this case we will always measure $\tilde{f}(\boldsymbol{\theta}) = 1$ [see Eq. (1)], so that $\langle \tilde{f}_A \rangle = \langle \tilde{f}_B \rangle = 1$, $\langle \tilde{f}_A \tilde{f}_B \rangle = 1$, and no error is expected on $\tilde{f}$. This result can be also recovered using the analytical expressions obtained so far. Let us first consider the simpler case of non-vanishing weights.

Using Eq. (37) and (38), we can write the term $T_{\mathrm{P}1}$ in the case $f(\boldsymbol{\theta}) = 1$ as

$$T_{\mathrm{P}1} = \rho \int_0^\infty \mathrm{d}s_A \int_0^\infty \mathrm{d}s_B \, \mathrm{e}^{\rho Q(s_A, s_B)} \times \int_\Omega \mathrm{d}^2\theta \, w_A(\boldsymbol{\theta}) w_B(\boldsymbol{\theta}) \mathrm{e}^{-s_A w_A(\boldsymbol{\theta}) - s_B w_B(\boldsymbol{\theta})} \, . \quad (59)$$

The second line in the r.h.s. of this equation can be rewritten as $\partial^2 Q / \partial s_A \, \partial s_B$ [cf. the definition of Q, Eq. (35)]:

$$T_{\mathrm{P}1} = \rho \int_0^\infty \mathrm{d}s_A \int_0^\infty \mathrm{d}s_B \, \mathrm{e}^{\rho Q(s_A, s_B)} \frac{\partial^2 Q(s_A, s_B)}{\partial s_A \partial s_B} \, . \quad (60)$$

Analogously, for $T_{\mathrm{P}2}$ we obtain [cf. Eq. (42)]

$$\begin{aligned} T_{\mathrm{P}2} &= \rho^2 \int_0^\infty \mathrm{d}s_A \int_0^\infty \mathrm{d}s_B \, \mathrm{e}^{\rho Q(s_A, s_B)} \\ &\quad \times \int_\Omega \mathrm{d}^2\theta_1 \, w_A(\boldsymbol{\theta}_1) \mathrm{e}^{-s_A w_A(\boldsymbol{\theta}_1) - s_B w_B(\boldsymbol{\theta}_1)} \\ &\quad \times \int_\Omega \mathrm{d}^2\theta_2 \, w_B(\boldsymbol{\theta}_2) \mathrm{e}^{-s_A w_A(\boldsymbol{\theta}_2) - s_B w_B(\boldsymbol{\theta}_2)} \\ &= \rho^2 \int_0^\infty \mathrm{d}s_A \int_0^\infty \mathrm{d}s_B \, \mathrm{e}^{\rho Q(s_A, s_B)} \frac{\partial Q(s_A, s_B)}{\partial s_A} \frac{\partial Q(s_A, s_B)}{\partial s_B} \, . \end{aligned} \quad (61)$$

We can integrate this expression by parts taking $\mathrm{e}^{\rho Q} \, (\partial Q / \partial s_B) = \bigl[\partial \exp(\rho Q) / \partial s_B \bigr] / \rho$ as differential term:

$$\begin{aligned} T_{\mathrm{P}2} = \rho \int_0^\infty \mathrm{d}s_A \biggl\{ & \biggl[\mathrm{e}^{\rho Q(s_A, s_B)} \frac{\partial Q(s_A, s_B)}{\partial s_A} \biggr]_{s_B = 0}^\infty \\ & - \int_0^\infty \mathrm{d}s_B \, \mathrm{e}^{\rho Q(s_A, s_B)} \frac{\partial^2 Q(s_A, s_B)}{\partial s_A \partial s_B} \biggr\} \, . \end{aligned} \quad (62)$$

We now observe that the last term in Eq. (62) is identical to what we founded in Eq. (60). Hence, the sum $T_{\mathrm{P}1} + T_{\mathrm{P}2}$ is

$$\begin{aligned} T_{\mathrm{P}1} + T_{\mathrm{P}2} &= \rho \biggl[\int_0^\infty \mathrm{d}s_A \, \mathrm{e}^{\rho Q(s_A, s_B)} \frac{\partial Q(s_A, s_B)}{\partial s_A} \biggr]_{s_B = 0}^\infty \\ &= \biggl[\Bigl[\mathrm{e}^{\rho Q(s_A, s_B)} \Bigr]_{s_A = 0}^\infty \biggr]_{s_B = 0}^\infty \\ &= Y(\infty, \infty) - Y(\infty, 0^+) - Y(0^+, \infty) + Y(0^+, 0^+) = 1 \, . \end{aligned} \quad (63)$$

The last equation holds because, for non-vanishing weights, $Y(0^+, 0^+) = 1$ and all other terms vanishes [cf. Eqs. (50–53)]. Hence, as expected, $\langle \tilde{f}_A \tilde{f}_B \rangle = T_{\mathrm{P}1} + T_{\mathrm{P}2} = 1 = \langle \tilde{f}_A \rangle \langle \tilde{f}_B \rangle$.

In case of vanishing weights, we can still use Eqs. (60) and (62) with an additional factor ν [due to the extra factor in Eq. (54)]. The last step in Eq. (63) thus now becomes

$$\begin{aligned} T_{\mathrm{P}1} + T_{\mathrm{P}2} = \nu \bigl[& Y(\infty, \infty) - Y(\infty, 0^+) \\ & - Y(0^+, \infty) + Y(0^+, 0^+) \bigr] = 1 \, . \end{aligned} \quad (64)$$

The last equality holds since now Y does not vanishes for large (s_A, s_B) [see again Eqs. (50–53)].

5.2. Scaling

Similarly to what was already shown in Paper I, for all expressions encountered so far some scaling invariance properties hold.

First, we note that, although we have assumed that the weight functions w_A and w_B are normalized to unity, all results are clearly independent of their actual normalization. Hence, a trivial scaling property holds: All results (and in particular the final expression for $\mathrm{Cov}(\tilde{f})$) are left unchanged by the transformation $w(\boldsymbol{\theta}) \mapsto k w(\boldsymbol{\theta})$ or, equivalently,

$$w_A(\boldsymbol{\theta}) \mapsto k w_A(\boldsymbol{\theta}) \; , \qquad\qquad w_B(\boldsymbol{\theta}) \mapsto k w_B(\boldsymbol{\theta}) \; . \tag{65}$$

A more interesting scaling property is the following. Consider the transformation

$$w(\boldsymbol{\theta}) \mapsto k^2 w(k\boldsymbol{\theta}) \; , \tag{66}$$

where the factor k^2 ensures that the transformed weight has the correct normalization (the exponent 2 in k^2 must be changed according to the dimension of the $\boldsymbol{\theta}$ vector space). If we apply this transformation, then the expression for $\mathrm{Cov}(\tilde{f})$ is transformed according to

$$\mathrm{Cov}(\tilde{f}; \boldsymbol{\theta}_A, \boldsymbol{\theta}_B) \mapsto \mathrm{Cov}(\tilde{f}; k\boldsymbol{\theta}_A, k\boldsymbol{\theta}_B) \; . \tag{67}$$

This invariance suggests that the shape of $\mathrm{Cov}(\tilde{f})$ is controlled by the expected number of objects for which the two weight functions are significantly different from zero. Hence, similarly to what done in Paper I, we define the two weight areas $\mathcal{A}_A$ and $\mathcal{A}_B$ as

$$\mathcal{A}_X \equiv \left[\int_\Omega \bigl[w_X(\boldsymbol{\theta}) \bigr]^2 \, \mathrm{d}^2\theta \right]^{-1} = \begin{cases} S_{20}^{-1} & \text{if } X = A \; , \\ S_{02}^{-1} & \text{if } X = B \; . \end{cases} \tag{68}$$

For weight functions invariant upon translation we have $\mathcal{A}_A = \mathcal{A}_B$. We call $\mathcal{N}_X \equiv \rho \mathcal{A}_X$ the *weight number of objects* (again, $\mathcal{N}_A = \mathcal{N}_B$ because of the invariance upon translation). Note that this quantity is left unchanged by the scaling (66). Similar definitions hold for the *effective weight* $w_{\mathrm{eff}\,X}(\boldsymbol{\theta}) \equiv w_X(\boldsymbol{\theta}) C_X\bigl(w_X(\boldsymbol{\theta})\bigr)$ and the *effective number of objects* $\mathcal{N}_{\mathrm{eff}\,X} \equiv \rho \mathcal{A}_{\mathrm{eff}\,X}$.

5.3. Behavior of C

In order to better understand the properties of C, it is useful to briefly consider its behavior as a function of the weights w_A and w_B.

We observe that, since $Y(s_A, s_B) > 0$ for every (s_A, s_B) [see Eq. (36)], $C(w_A, w_B)$ decreases if either w_A or w_B increase. In order to study the behavior of the quantity $w_A w_B C(w_A, w_B)$ that enters the noise term T_1, we consider the quantity $w_A C(w_A, w_B)$:

$$w_A C(w_A, w_B) = \nu\rho^2 \int_{0^-}^{\infty} \mathrm{d}s_B \Biggl[Y(0^-, s_B) + \int_{0^-}^{\infty} \mathrm{d}s_A \left(\frac{\partial Y(s_a, s_B)}{\partial s_A} \right) \mathrm{e}^{-s_A w_A} \Biggr] \mathrm{e}^{-s_B w_B} \; . \tag{69}$$

This equation can be shown by integrating by parts the integral over s_A. The partial derivative required in Eq. (69) can be evaluated from Eq. (36):

$$\frac{\partial Y(s_A, s_B)}{\partial s_A} = \rho \frac{\partial Q(s_A, s_B)}{\partial s_A} \mathrm{e}^{\rho Q(s_A, s_B)} \le 0 \; . \tag{70}$$

Since this derivative is negative, we can deduce that the integral over s_A in Eq. (69) increases with w_A, and thus $w_A C(w_A, w_B)$ also increases as w_A increases. Similarly,

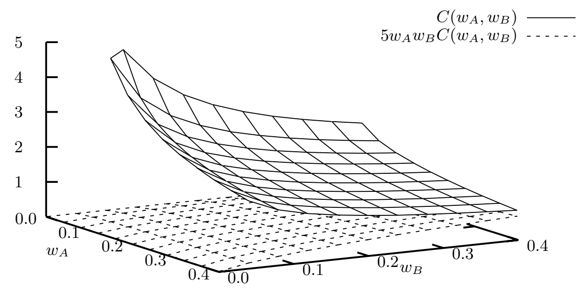

Fig. 2. The function $C(w_A, w_B)$ is monotonically decreasing with w_A and w_B, while $w_A w_B C(w_A, w_B)$ (scaled in this plot) is monotonically increasing. The parameters used for this figure are the same as Fig. 1. Note that, since $P_A = P_B = 0$, we have $\lim_{w_A \to 0^+} w_A w_B C(w_A, w_B) = \lim_{w_B \to 0^+} w_A w_B C(w_A, w_B) = 0$ in agreement with Eqs. (72) and (73); moreover $w_A w_B C(w_A, w_B) < \rho^2 = 1$ as expected from Eq. (74).

it can be shown that $w_B C(w_A, w_B)$ increases as w_B increases. In summary, the quantity $w_A w_B C(w_A, w_B)$ behaves as $w_A w_B$, in the sense that its partial derivatives have the same sign as the partial derivatives of $w_A w_B$ (see Fig. 2). Also, since $C(w_A, w_B)$ decreases if either w_A or w_B increase, we can deduce that $w_A w_B C(w_A, w_B)$ is "broader" than $w_A w_B$.

Since $C\bigl(w_A(\boldsymbol{\theta}), w_B(\boldsymbol{\theta})\bigr)$ is positive, the function $w_A(\boldsymbol{\theta}) w_B(\boldsymbol{\theta}) C\bigl(w_A(\boldsymbol{\theta}), w_B(\boldsymbol{\theta})\bigr)$ shares the same support as $w_A(\boldsymbol{\theta}) w_B(\boldsymbol{\theta})$. It is also interesting to study the limits of $w_A w_B C(w_A, w_B)$ at high and low values for w_A and w_B. From the properties of Laplace transform [see Eqs. (C.10) and (C.11)], we have

$$\lim_{\substack{w_A \to 0^+ \\ w_B \to 0^+}} w_A w_B C(w_A, w_B) = \nu\rho^2 \lim_{\substack{s_A \to \infty \\ s_B \to \infty}} Y(s_A, s_B) = \nu\rho^2 P_{AB} \; , \tag{71}$$

where Eq. (50) has been used in the second equality. Hence, the quantity $w_A w_B C(w_A, w_B)$ goes to zero only if $P_{AB} = 0$. In other cases, we expect a discontinuity at $w_A = w_B = 0$. Similarly, using Eqs. (50–52) we find

$$\lim_{\substack{w_A \to \infty \\ w_B \to 0^+}} w_A w_B C(w_A, w_B) = \nu\rho^2 \lim_{\substack{s_A \to 0^+ \\ s_B \to \infty}} Y(s_A, s_B) = \nu\rho^2 P_A \; . \tag{72}$$

$$\lim_{\substack{w_A \to 0^+ \\ w_B \to \infty}} w_A w_B C(w_A, w_B) = \nu\rho^2 \lim_{\substack{s_A \to \infty \\ s_B \to 0^+}} Y(s_A, s_B) = \nu\rho^2 P_B \; . \tag{73}$$

$$\lim_{\substack{w_A \to \infty \\ w_B \to \infty}} w_A w_B C(w_A, w_B) = \nu\rho^2 \lim_{\substack{s_A \to 0^+ \\ s_B \to 0^+}} Y(s_A, s_B) = \nu\rho^2 \; . \tag{74}$$

Since $w_A w_B C(w_A, w_B)$ increases with both w_A and w_B, the last equation above puts a superior limit for this quantity:

$$w_A w_B C(w_A, w_B) \le \nu \rho^2 . \tag{75}$$

5.4. Distant regions

Suppose that the two points $\boldsymbol\theta_A$ and $\boldsymbol\theta_B$ are far away from each other, so that $w_A(\boldsymbol\theta) w_B(\boldsymbol\theta)$ is very close to zero everywhere. In this situation we can greatly simplify our equations.

If $\boldsymbol\theta_A$ is far away from $\boldsymbol\theta_B$, then $w_A(\boldsymbol\theta)$ and $w_B(\boldsymbol\theta)$ are never significantly different from zero at the same position $\boldsymbol\theta$. In this case, the integral in the definition of $Q(s_A, s_B)$ [see Eq. (35)] can be split into two integrals that corresponds to Q_A and Q_B [Eq. (43)]:

$$Q(s_A, s_B) \simeq Q_A(s_A) + Q_B(s_B) , \tag{76}$$

$$Y(s_A, s_B) \simeq Y_A(s_A) Y_B(s_B) , \tag{77}$$

$$C(w_A, w_B) \simeq C_A(w_A) C_B(w_B) . \tag{78}$$

Hence, if the two weight functions w_A and w_B do not have significant overlap, the function $C(w_A, w_B)$ reduces to the product of the two correcting functions C_A and C_B.

In general, it can be shown that $C(w_A, w_B) \ge C_A(w_A) C_B(w_B)$. In fact, we have

$$C(w_A, w_B) - C_A(w_A) C_B(w_B) = \rho^2 \int_0^\infty \mathrm{d}s_A \int_0^\infty \mathrm{d}s_B \, \mathrm{e}^{-s_A w_A - s_B w_B} \times \left[\mathrm{e}^{\rho Q(s_A, s_B)} - \mathrm{e}^{\rho Q_A(s_A) + \rho Q_B(s_B)} \right] . \tag{79}$$

We now observe that

$$Q(s_A, s_B) - Q_A(s_A) - Q_B(s_B) = \int_\Omega \left[\mathrm{e}^{-s_A w_A(\boldsymbol\theta)} - 1 \right] \left[\mathrm{e}^{-s_B w_B(\boldsymbol\theta)} - 1 \right] \ge 0 . \tag{80}$$

Hence, $Q(s_A, s_B) \ge Q_A(s_A) + Q_B(s_B)$ and the difference between the two terms of this inequality is an indication of overlap between the two weight functions w_A and w_B. Since the exponential function is monotonic, we find $Y(s_A, s_B) \ge Y_A(s_A) Y_B(s_B)$ and thus

$$C(w_A, w_B) \ge C_A(w_A) C_B(w_B) . \tag{81}$$

5.5. Upper and lower limits for T_σ

The normalization property shown in Sect. 5.1 can also be used to obtain an upper limit for T_σ. We observe, in fact, that T_σ is indistinguishable from $\sigma^2 T_{\mathrm{P}1}$ for a constant function $f(\boldsymbol\theta) = 1$. This case, however, has already been considered above in Sect. 5.1: There we have shown that $T_{\mathrm{P}1} + T_{\mathrm{P}2} = 1$. Since $T_{\mathrm{P}2} \ge 0$, we find the relation $T_\sigma \le \sigma^2$.

The property just obtained has a simple interpretation. As shown by Eq. (49), T_σ is proportional to $1/\rho$ and thus we would expect that this quantity is unbounded superiorly. In reality, even when we are dealing with a very small density of objects, the estimator (1) "forces" us to use at least one object. This point has already been discussed in Paper I, where we showed that the number of effective objects, $\mathcal{N}_{\mathrm{eff}}$, is always larger than unity. The upper limit found for T_σ can be interpreted using the same argument. Note that this result also holds for weight functions with finite support.

A lower limit for T_σ, instead, can be obtained from the inequality (81):

$$T_\sigma \ge \frac{\sigma^2}{\rho} \int_\Omega w_A(\boldsymbol\theta) w_B(\boldsymbol\theta) C_A\bigl(w_A(\boldsymbol\theta)\bigr) C_B\bigl(w_B(\boldsymbol\theta)\bigr) \, \mathrm{d}^2\theta = \frac{\sigma^2}{\rho} \int_\Omega w_{\mathrm{eff}A}(\boldsymbol\theta) w_{\mathrm{eff}B}(\boldsymbol\theta) \, \mathrm{d}^2\theta . \tag{82}$$

Hence, the error T_σ is larger than a convolution of the two effective weight functions. In case of finite-field weight functions, the limit just obtained must be corrected with a factor ν. The argument to derive Eq. (82) is then slightly more complicated because of the presence of the P_X probabilities. However, using the relation $P_A P_B \ge P_{AB}$, we can recover Eq. (82) with the aforementioned corrective factor.

5.6. Limit of low and high densities

In the limit $\rho \to 0$ we can obtain simple expressions for the noise terms. If ρ vanishes, we have $Y(s_A, s_B) = 1$ [cf. Eq. (36)] and thus

$$C(w_A, w_B) \simeq \frac{\nu \rho^2}{w_A w_B} . \tag{83}$$

In this equation we have assumed $w_A w_B > 0$. Note that we have reached here the superior limit indicated by Eq. (75). In the same limit, $\rho \to 0$, $P_X \simeq 1 - \pi_X \rho$ and $\nu \simeq 1/\rho \pi_{A\cap B}$, where $\pi_{A\cap B} = \pi_A + \pi_B - \pi_{A\cup B}$ is the area of the intersection of the supports of w_A and w_B. Hence we find

$$C(w_A, w_B) \simeq \frac{\rho}{\pi_{A\cap B} w_A w_B} . \tag{84}$$

Analogously, in the same limit, we have found in Paper I

$$C_X(w_X) \simeq \frac{1}{\pi_X w_X} , \tag{85}$$

where $w_X > 0$ has been assumed. We can then proceed to evaluate the various terms. For T_σ we obtain the expression

$$T_\sigma \simeq \frac{\sigma^2}{\rho} \int_{\pi_{A\cap B}} \frac{\rho}{\pi_{A\cap B}} \, \mathrm{d}^2\theta = \sigma^2 . \tag{86}$$

Note that the integral has been evaluated only on the subset of the plane where $w_A w_B > 0$; the case where this product vanishes, in fact, need not to be considered because the quantity $w_A w_B C(w_A, w_B)$ vanishes as well. Exactly the same result holds for weight functions with infinite support. Hence, when $\rho \to 0$ we reach the superior limit discussed in Sect. 5.5 for T_σ.

Equation (86) can be better appreciated with the following argument. As the density ρ approaches zero, the probability of having two objects on $\pi_{A\cup B}$ vanishes. Because of the prescription regarding vanishing weights (cf. beginning of Sect. 3), the ensemble average in our limit is performed with one and only one object in $\pi_{A\cap B}$. Since we have only one object, this is basically used with unit weight in the average (8), and thus the measurement noise is just given by $T_\sigma = \sigma^2$.

Let us now consider the limit at low densities of the Poisson noise, which, we recall, has been split into three terms, T_{P1}, T_{P2}, and T_{P3} (see Sect. 2.3). Inserting Eq. (84) into Eq. (15), we find for T_{P1}

$$T_{\mathrm{P1}} \simeq \frac{1}{\rho} \int_{\pi_{A\cap B}} f^2(\boldsymbol{\theta}) \frac{\rho}{\pi_{A\cap B}} \,\mathrm{d}^2\theta = \bigl\langle f^2 \bigr\rangle_{\pi_{A\cap B}} , \tag{87}$$

where $\bigl\langle f^2 \bigr\rangle_{\pi_{A\cap B}}$ denotes the simple average of f^2 on the set $\pi_{A\cap B}$. Hence, T_{P1} converges to the average of f^2 on the intersection of the supports of w_A and w_B. Again, we can explain this result using an argument similar to the one used for Eq. (86). Regarding $T_{\mathrm{P2}} \equiv T_2$, we observe that this term is of first order in ρ because $C(w_A, w_B)$ is of first order [cf. Eqs. (84) and (42)]. We can then safely ignore this term in our limit $\rho \to 0$. Finally, as shown in Paper I, at low densities the expectation value for $\tilde{f}_X$ is a simple average of f on the support of w_X, i.e. $\langle \tilde{f}_X \rangle \simeq \langle f \rangle_{\pi_X}$. Hence, $T_{\mathrm{P3}} = \langle f \rangle_{\pi_A} \langle f \rangle_{\pi_B}$ and the Poisson noise in the limit of small densities is given by

$$T_{\mathrm{P}} = \bigl\langle f^2 \bigr\rangle_{\pi_{A\cap B}} - \langle f \rangle_{\pi_A} \langle f \rangle_{\pi_B} . \tag{88}$$

In case of a constant function $f(\boldsymbol{\theta})$, this expression vanishes as expected. Surprisingly, in general, we cannot say that $T_{\mathrm{P}} \ge 0$. Rather, if $w_A \neq w_B$, and if in particular the two weight functions have different supports, we might have a negative T_{P}. Suppose, for example, that f vanishes on the intersection of the two supports $\pi_{A\cap B}$, but is otherwise positive. In this case, the first term in the r.h.s. of Eq. (88) vanishes, while the second term contributes with a negative sign, and thus $T_{\mathrm{P}} < 0$. On the other hand, if $w_A = w_B$ then T_{P} has to be non-negative.

We now consider the opposite limiting case, namely high density. In this case, it is useful to use the moment expansion (58). Since T_σ and T_{P1} have an overall factor $1/\rho$ in its definition [cf. Eq. (49)], we can simply take the 0-th order for $C(w_A, w_B)$, thus obtaining

$$T_\sigma \simeq \frac{\sigma^2}{\rho} \int w_A(\boldsymbol{\theta}) w_B(\boldsymbol{\theta}) \,\mathrm{d}^2\theta = \frac{\sigma^2 S_{11}}{\rho} . \tag{89}$$

Similarly, for T_{P1} we obtain

$$T_{\mathrm{P1}} \simeq \frac{1}{\rho} \int w_A(\boldsymbol{\theta}) w_B(\boldsymbol{\theta}) f^2(\boldsymbol{\theta}) \,\mathrm{d}^2\theta . \tag{90}$$

For T_{P2} and T_{P3}, instead, we need to use a first order expansion in $1/\rho$ for $C(w_A, w_B)$. This can be done by using the first terms in series (56), and by expanding all fractions in terms of powers of $1/\rho$. Inserting the result into the definitions of T_{P2} and T_{P3} we obtain

$$\begin{aligned} T_{\mathrm{P2}} \simeq{}& \int_\Omega \mathrm{d}^2\theta_1 \int_\Omega \mathrm{d}^2\theta_2 \, f(\boldsymbol{\theta}_1) f(\boldsymbol{\theta}_2) w_A(\boldsymbol{\theta}_1) w_B(\boldsymbol{\theta}_2) \\ &\times \left[1 - \frac{w_A(\boldsymbol{\theta}_1)}{\rho} - \frac{w_A(\boldsymbol{\theta}_2)}{\rho} - \frac{w_B(\boldsymbol{\theta}_1)}{\rho} - \frac{w_B(\boldsymbol{\theta}_2)}{\rho} \right. \\ &\left. + \frac{S_{20} + S_{11} + S_{02}}{\rho} \right] , \end{aligned} \tag{91}$$

$$\begin{aligned} T_{\mathrm{P3}} \simeq{}& \int_\Omega \mathrm{d}^2\theta_1 \int_\Omega \mathrm{d}^2\theta_2 \, f(\boldsymbol{\theta}_1) f(\boldsymbol{\theta}_2) w_A(\boldsymbol{\theta}_1) w_B(\boldsymbol{\theta}_2) \\ &\times \left[1 - \frac{w_A(\boldsymbol{\theta}_1)}{\rho} - \frac{w_B(\boldsymbol{\theta}_2)}{\rho} + \frac{S_{20} + S_{02}}{\rho} \right] . \end{aligned} \tag{92}$$

Note that we have dropped, in these equations, terms of order higher than $1/\rho$. The difference $T_{\mathrm{P2}} - T_{\mathrm{P3}}$ is

$$\begin{aligned} T_{\mathrm{P2}} - T_{\mathrm{P3}} \simeq{}& \frac{1}{\rho} \int_\Omega \mathrm{d}^2\theta_1 \int_\Omega \mathrm{d}^2\theta_2 \, f(\boldsymbol{\theta}_1) f(\boldsymbol{\theta}_2) w_A(\boldsymbol{\theta}_1) w_B(\boldsymbol{\theta}_2) \\ &\times \bigl[S_{11} - w_A(\boldsymbol{\theta}_2) - w_B(\boldsymbol{\theta}_1) \bigr] . \end{aligned} \tag{93}$$

Using Eqs. (93) and (90), we can verify that T_{P} vanishes if f is constant, as expected:

$$\begin{aligned} T_{\mathrm{P1}} + T_{\mathrm{P2}} - T_{\mathrm{P3}} \simeq{}& \frac{S_{11}}{\rho} + \frac{1}{\rho} \int_\Omega \mathrm{d}\theta_1 \int_\Omega \mathrm{d}\theta_2 \, \bigl[S_{11} w_A(\theta_1) w_B(\theta_2) \\ &- w_A(\theta_1) w_A(\theta_2) w_B(\theta_2) \\ &- w_A(\theta_1) w_B(\theta_2) w_B(\theta_1) \bigr] \\ ={}& \frac{S_{11}}{\rho} + \frac{1}{\rho} \bigl[S_{11} - S_{11} - S_{11} \bigr] = 0 , \end{aligned} \tag{94}$$

where the normalization of w has been used. Also, it is apparent that all noise sources, including Poisson noise, are proportional to $1/\rho$ at high densities.

In order to further investigate the properties of Poisson noise at high densities, we write it in a more compact form. Let us define the average of a function $g(\boldsymbol{\theta})$ weighted with $q(\boldsymbol{\theta})$ as

$$\langle g \rangle_q \equiv \left[\int_\Omega g(\boldsymbol{\theta}) q(\boldsymbol{\theta}) \,\mathrm{d}^2\theta \right] \bigg/ \left[\int_\Omega q(\boldsymbol{\theta}) \,\mathrm{d}^2\theta \right] . \tag{95}$$

Using this definition we can rearrange Eqs. (90) and (93) in the form

$$\begin{aligned} T_{\mathrm{P}} ={}& \frac{S_{11}}{\rho} \bigl[\bigl\langle f^2 \bigr\rangle_{w_A w_B} - \langle f \rangle_{w_A w_B} \langle f \rangle_{w_A w_B} \\ &+ \bigl(\langle f \rangle_{w_A w_B} - \langle f \rangle_{w_A} \bigr) \bigl(\langle f \rangle_{w_A w_B} - \langle f \rangle_{w_B} \bigr) \bigr] . \end{aligned} \tag{96}$$

This expression suggests that the Poisson noise is actually made of two different terms, represented by the first and the second lines of Eq. (96). The first term is proportional to the difference between two averages of f^2 and f; both averages are performed using $w_A w_B$ as weight. Hence, this term is controlled by the "internal scatter" of f on points where both weight functions are significantly different from zero; it is always positive. The second term is made of averages f using different weight functions. It can be either positive or negative if $w_A \neq w_B$. Actually, as already seen in the limiting case $\rho \to 0$, the overall Poisson noise does not need to be positive, and anti-correlation can be present in some cases.

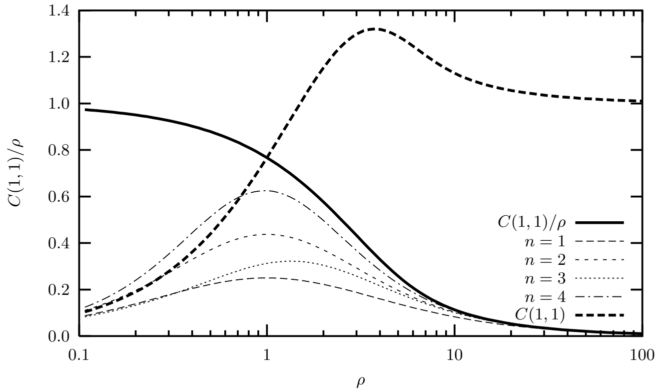

Fig. 3. The value of $C(1,1)/\rho$ for top-hat weights as a function of the density ρ. Both weight functions w_A and w_B are top-hats [see Eq. (97)] centered on zero.

6. Examples

Similarly to what has been done in Paper I, in this section we consider three typical weight functions, namely a top-hat, a Gaussian, and a parabolic weight. For simplicity, we will consider 1-dimensional cases only; this will have also some advantages when representing the results obtained with figures. Hence, we will use x instead of $\boldsymbol{\theta}$ as spatial variable.

6.1. Top-hat

The simplest weight that we can consider is a top-hat function, defined as

$$w(x) = \begin{cases} 1 & \text{if } |x| < 1/2 \; , \\ 0 & \text{otherwise} \; . \end{cases} \tag{97}$$

Since w is either 1 or 0, we just need to consider $C(1,1)$ to evaluate T_σ. Regarding the Poisson noise, from Eq. (42) we deduce that $C(1,2)$, $C(2,1)$, and $C(2,2)$ are also required.

Figure 3 shows $C(1,1)$ and $C(1,1)/\rho$ as functions of the density ρ for two identical top-hat weight functions centered on the origin. From this plot we can recognize some of the limiting cases studied above. In particular, the fact that $C(1,1)/\rho$ goes to unity at low densities is related to Eq. (84); similarly, the limit of $C(1,1)$ at high densities is consistent with Eq. (58). The same figure shows also the moments expansion of $C(1,1)$ up to forth order. As expected, the expansion completely fails at low densities, while is quite accurate for $\rho > 5$.

Curves in Fig. 3 have been calculated using the standard approach described by Eqs. (35), (36) and (54). Actually, in the simple case of top-hat weight functions, we can evaluate $C(1,1)$ using a more direct statistical argument. We start by observing that in our case we have $T_\sigma = \sigma^2 C(1,1)/\rho$. On the other hand, a top-hat weight function is basically acting by taking simple averages for all objects that fall inside its support. This suggests that we can evaluate its measurement noise as

$$T_\sigma = \sigma^2 \sum_{N=1}^{\infty} \frac{p(N)}{N} \; , \tag{98}$$

where $p(N)$ is the probability of having N objects inside the support. This probability is basically a Poisson probability distribution with average ρ. However, since we are adopting the prescription of "avoiding" weight functions without objects in their support, we must explicitely discard the case $N = 0$ and consequently renormalize the probability. In summary, we have

$$p(N) = \frac{\mathrm{e}^{-\rho} \rho^N}{N!} \Big/ \left[1 - \mathrm{e}^{-\rho} \right] \; . \tag{99}$$

This expression combined with Eq. (98) allows us to evaluate $C(1,1) = \rho T_\sigma / \sigma^2$:

$$C(1,1) = \frac{\mathrm{e}^{-\rho}}{1 - \mathrm{e}^{-\rho}} \sum_{N=1}^{\infty} \frac{\rho^{N+1}}{N! \, N} \; . \tag{100}$$

We can directly verify this result using Eqs. (35), (36) and (54). In fact, for the top-hat function we find

$$Q(s_A, s_B) = \mathrm{e}^{-s_A - s_B} - 1 \; , \tag{101}$$

$$Y(s_A, s_B) = \mathrm{e}^{\rho Q(s_A, s_B)} = \mathrm{e}^{-\rho} \sum_{k=0}^{\infty} \frac{\mathrm{e}^{-k(s_A + s_B)} \rho^k}{k!} \; , \tag{102}$$

$$\begin{aligned} C(1,1) &= \frac{\rho^2}{1 - \mathrm{e}^{-\rho}} \int_0^\infty \mathrm{d}s_A \int_0^\infty \mathrm{d}s_B Y(s_A, s_B) \mathrm{e}^{-s_A - s_B} \\ &= \frac{\rho^2 \mathrm{e}^{-\rho}}{1 - \mathrm{e}^{-\rho}} \sum_{k=0}^{\infty} \int_0^\infty \mathrm{d}s_A \int_0^\infty \mathrm{d}s_B \frac{\mathrm{e}^{-(k+1)(s_A + s_B)} \rho^k}{k!} \\ &= \frac{\mathrm{e}^{-\rho}}{1 - \mathrm{e}^{-\rho}} \sum_{k=0}^{\infty} \frac{\rho^{k+2}}{k! \, (k+1)^2} \; . \end{aligned} \tag{103}$$

Finally, with a change of the dummy variable $k \mapsto n - 1$ we recover Eq. (100).

The other terms needed for the Poisson noise can be evaluated using a calculation similar to the one performed in Eq. (103). Actually, it can be shown that for any positive integers w_A and w_B we have

$$C(w_A, w_B) = \frac{\mathrm{e}^{-\rho}}{1 - \mathrm{e}^{-\rho}} \sum_{k=0}^{\infty} \frac{\rho^{k+2}}{k! \, (k + w_A)(k + w_B)} \; . \tag{104}$$

6.2. Gaussian

Frequently, a Gaussian weight function of the form

$$w(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2} \tag{105}$$

is used. Although it is not possible to carry out analytical calculations and obtain $C(w_A, w_B)$, numerical integrations do not pose any problem. Figure 4 shows, for different densities, the function $w_A w_B C(w_A, w_B)$ for two identical weights $w_A = w_B$ centered in zero; Fig. 5 shows the same quantity when one of the weight function is centered at unity. Note that, in this last figure, the largest covariance is at $x = 0.5$, as expected.

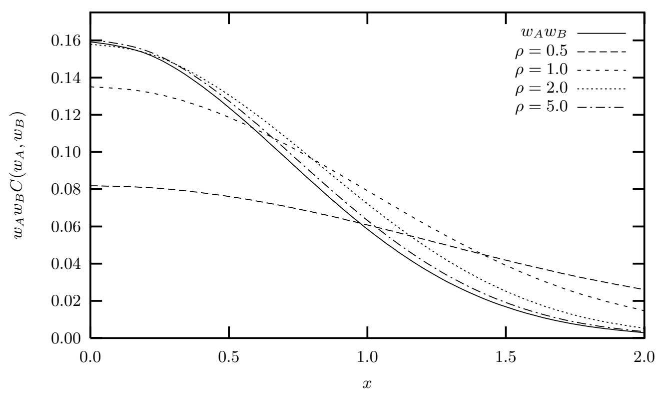

Fig. 4. Numerical calculations for 1-dimensional Gaussian weight functions $w_A = w_B$ centered on 0 and with unit variance. The various curves shows the function $w_A w_B C(w_A, w_B)$ for different densities ρ. Note that, as expected, $C(w_A, w_B)$ approaches unity for large densities.

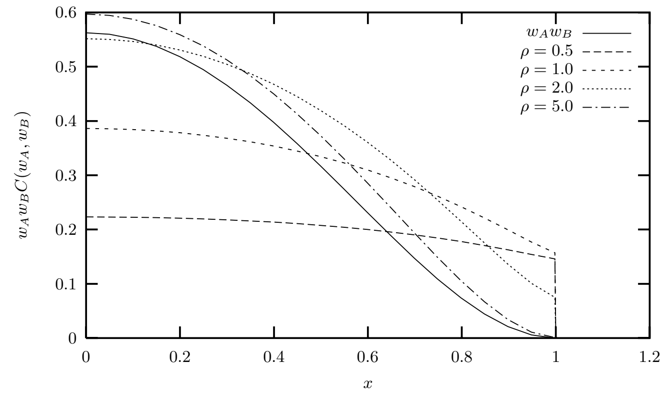

Fig. 6. Numerical calculations for 1-dimensional parabolic weight functions $w_A = w_B$ centered on 0 and with unit variance. The various curves shows the function $w_A w_B C(w_A, w_B)$ for different densities ρ.

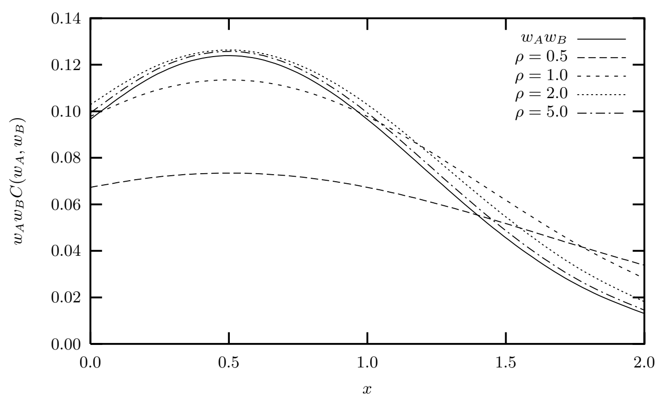

Fig. 5. Same as Fig. 4, but for two Gaussian weight functions centered on 0 and 1 and with unit variance.

6.3. Parabolic weight

Finally, we study of a parabolic weight function of the form

$$w(x) = \begin{cases} 3x^2/4 & \text{if } |x| < 1 \,, \\ 0 & \text{otherwise} \,. \end{cases} \tag{106}$$

This function illustrates well some of the peculiarities of finite support weight functions. Figure 6 shows the results of numerical integrations for $w_A w_B C(w_A, w_B)$ at different densities ρ. A first interesting point to note is the discontinuity observed at $x = 1$, which is in agreement with Eq. (71). Moreover, as expected from Eq. (84), the function plotted clearly approaches a constant at low densities ρ.

7. Conclusions

In this article we have studied in detail the covariance of a widely used smoothing technique. The main results obtained are summarized in the following items.

1. The covariance is composed of two main terms, T_σ and T_P, representing measurement errors and Poisson noise, respectively.
2. Expressions to compute T_σ and T_P have been provided. In particular, it has been shown that both terms can be obtained in term of a kernel $C(w_A, w_B)$, which in turn can be evaluated from the weight function $w(\boldsymbol{\theta})$.
3. We have obtained an expansion of the kernel $C(w_A, w_B)$ valid at high densities ρ.
4. We have shown that T_σ has an upper limit, given by σ^2, and a lower limit, provided by Eq. (82).
5. We have evaluated the form of the noise contributions in the limiting cases of high and low densities.
6. We have considered three typical cases of weight functions and we have evaluated $C(w_A, w_B)$ for them.

Finally, we note that although the smoothing technique considered in this paper is by far the most widely used in Astronomy, alternative methods are available. A statistical characterization of these methods, using a completely different approach, will be presented in a future paper (Lombardi & Schneider, in preparation).

Appendix A: Vanishing weights

In Sect. 2.2 we have obtained the solution of the covariance problem under the hypothesis that the weight function $w(\boldsymbol{\theta})$ is strictly positive. In this appendix we will generalize the results obtained there to non-negative weight functions (see also Sect. 3).

If w_A is allowed to vanish, then we might have a finite probability that y_A vanishes, i.e. a finite probability that no point $\boldsymbol{\theta}_n$ is inside the support of w_A. A finite probability in a probability distribution function appears as a Dirac's delta distribution. Since this point is quite important for our discussion, let us make a simple example. Suppose that ξ is a real random variable with the following characteristics:

– ξ has probability 1/3 to vanish.

– ξ has probability 2/3 to be in the range $(0, \infty)$; in this range ξ has an exponential distribution.

Then we can write the probability distribution function for ξ as

$$p_\xi(\xi) = \frac{1}{3}\delta(\xi) + \frac{2}{3}\mathrm{H}(\xi)\exp(-\xi) , \tag{A.1}$$

where H is the Heaviside function [see Eq. (30)]. In other words, the probability distribution for ξ includes the contribution from a Dirac's delta distribution centered on $\xi = 0$. If p_ξ is known, the probability that ξ is exactly zero (1/3 in this example) can be obtained using

$$P(\xi = 0) = \int_{0^-}^{0^+} p_\xi(\xi')\,\mathrm{d}\xi' = \lim_{\xi \to 0^+} \int_0^\xi p_\xi(\xi')\,\mathrm{d}\xi' . \tag{A.2}$$

Let us now turn to our problem. As mentioned above, for vanishing weights we expect that y_A might vanish, i.e. its probability might include the contribution from a delta distribution centered on $y_A = 0$; similarly, if w_B is allowed to vanish, the probability distribution for y_B might include a delta centered in $y_B = 0$. For a given y_B, the probability $P_A(y_B)$ that y_A vanishes is given by

$$P_A(y_B) \equiv \lim_{y_A \to 0^+} \int_{0^-}^{y_A} p_y(y'_A, y_B)\,\mathrm{d}y'_A = \lim_{s_A \to \infty} \mathcal{L}_A\bigl[p_y(\cdot, y_B)\bigr](s_A) \tag{A.3}$$

where the properties of Laplace transform have been used in the last equality (see Appendix C). A similar equation holds for the probability that y_B vanishes, $P_B(y_A)$. Note that the Laplace transform in Eq. (A.3) is performed only with respect to the first variable. The joint probability P_{AB} that both y_A and y_B vanish is [cf. Eq. (50)]

$$\begin{aligned} P_{AB} &\equiv \lim_{\substack{y_A \to 0^+ \\ y_B \to 0^+}} \int_{0^-}^{y_A} \mathrm{d}y'_A \int_{0^-}^{y_B} \mathrm{d}y'_B\, p_y(y'_A, y'_B) \\ &= \lim_{\substack{s_A \to \infty \\ s_B \to \infty}} \mathcal{L}[p_y](s_A, s_B) = Y(\infty, \infty) . \end{aligned} \tag{A.4}$$

We then also define [cf. Eqs. (51) and (52)]

$$P_A \equiv \int_0^\infty P_A(y_B)\,\mathrm{d}y_B = \mathcal{L}[p_y](\infty, 0^+) = Y(\infty, 0^+) . \tag{A.5}$$

$$P_B \equiv \int_0^\infty P_B(y_A)\,\mathrm{d}y_A = \mathcal{L}[p_y](0^+, \infty) = Y(0^+, \infty) . \tag{A.6}$$

Using Eq. (35), we find $P_A = \exp(-\rho\pi_A)$, $P_B = \exp(-\rho\pi_B)$, and $P_{AB} = \exp(-\rho\pi_{A\cap B})$, where π_A is the area of the support of w_A, π_B is the area of the support of w_B, and $\pi_{A\cap B}$ is the area of the intersection of the two supports. This result is of course not surprising and has been already derived in the paragraph before Eq. (50) using a different approach.

For vanishing weights, we decided to use the following prescription: We discard, in the ensemble average for $\mathrm{Cov}(\tilde{f}; \boldsymbol{\theta}_A, \boldsymbol{\theta}_B)$, the configurations $\{\boldsymbol{\theta}_n\}$ for which the function $\tilde{f}$ is not defined either at $\boldsymbol{\theta}_A$ or at $\boldsymbol{\theta}_B$. In order to implement this prescription, we can explicitly modify the probability distribution p_y and exclude "by hand" cases where the denominator of Eq. (10) vanishes; for the purpose, we consider separately cases where w_A or w_B vanish. We define a new probability distribution for (y_A, y_B) which accounts for vanishing weights:

$$\bar{p}_y(y_A, y_B) \equiv \begin{cases} p_y(y_A, y_B) & \text{if } w_A \neq 0,\ w_B \neq 0 , \\ \bigl[p_y(y_A, y_B) - P_A(y_B)\delta(y_A)\bigr] & \\ \quad /(1 - P_A) & \text{if } w_A = 0,\ w_B \neq 0 , \\ \bigl[p_y(y_A, y_B) - P_B(y_A)\delta(y_B)\bigr] & \\ \quad /(1 - P_B) & \text{if } w_A \neq 0,\ w_B = 0 , \\ \bigl[p_y(y_A, y_B) - P_A(y_B)\delta(y_A) & \\ \qquad - P_B(y_A)\delta(y_B) & \\ \qquad + P_{AB}\delta(y_A)\delta(y_B)\bigr] & \\ \quad /(1 - P_B - P_B + P_{AB}) & \text{if } w_A = 0,\ w_B = 0 . \end{cases} \tag{A.7}$$

In constructing this probability, first we have explicitly removed the degenerate situations, then we have renormalized the resulting probability. Note that the normalization factor in the last case, namely $1 - P_A - P_B + P_{AB}$, comes from the so-called "inclusion-exclusion principle." Using this new probability distribution in the definition (23) for Y we obtain

$$Y(s_A, s_B) = \begin{cases} \mathrm{e}^{\rho Q(s_A, s_B)} & \text{if } w_A \neq 0,\ w_B \neq 0 , \\ \bigl[\mathrm{e}^{\rho Q(s_A, s_B)} - \mathrm{e}^{\rho Q(\infty, s_B)}\bigr] & \\ \quad /(1 - P_A) & \text{if } w_A = 0,\ w_B \neq 0 , \\ \bigl[\mathrm{e}^{\rho Q(s_A, s_B)} - \mathrm{e}^{\rho Q(s_A, \infty)}\bigr] & \\ \quad /(1 - P_B) & \text{if } w_A \neq 0,\ w_B = 0 , \\ \bigl[\mathrm{e}^{\rho Q(s_A, s_B)} - \mathrm{e}^{\rho Q(s_A, \infty)} & \\ \qquad - \mathrm{e}^{\rho Q(\infty, s_B)} + \mathrm{e}^{\rho Q(\infty, \infty)}\bigr] & \\ \quad /(1 - P_A - P_B + P_{AB}) & \text{if } w_A = 0,\ w_B = 0 . \end{cases} \tag{A.8}$$

Finally, we need to change the normalization factor in Eq. (37) in order to account for cases where y_A or y_B are vanishing:

$$C(w_A, w_B) = \frac{\rho^2}{1 - P_A - P_B + P_{AB}} \mathcal{L}[Y](w_A, w_B) . \tag{A.9}$$

This complete the discussion of vanishing weights.

Appendix B: Moments expansion

In Sect. 4 we have written the moments expansion for $C(w_A, w_B)$. Here we complete the discussion by providing a proof for that result.

At high densities, y_A and y_B are basically Gaussian random variables with average values $\bar{y}_A$ and $\bar{y}_B$ (we anticipate here that these averages are given by the density ρ). Hence, we can expand them in the definition of

Order $(i+j)$	M_{ij} $j=0$	$j=1$	$j=2$	$j=3$	$j=4$
0	1	–	–	–	–
1	0	0	–	–	–
2	ρS_{20}	ρS_{11}	ρS_{02}	–	–
3	ρS_{30}	ρS_{21}	ρS_{12}	ρS_{03}	–
4	$\rho S_{40} + 3\rho^2 S_{20}$	$\rho S_{30} + 3\rho^2 S_{20} S_{11}$	$\rho S_{22} + \rho^2 S_{02} S_{20} + 2\rho^2 S_{11} S_{11}$	$\rho S_{03} + 3\rho^2 S_{02} S_{11}$	$\rho S_{04} + 3\rho^2 S_{02}$

Table B.1. Moments M_{ij} up to the fourth order. The table shows, for each row, the values of M_{ij} with $(i+j)$, the order, fixed. Hence, for example, the row for order 2 shows M_{20}, M_{11}, and M_{02} in sequence.

$C(w_A, w_B)$:

$$\begin{aligned} C(w_A, w_B) &= \rho^2 \int_0^\infty \mathrm{d}y_A \int_0^\infty \mathrm{d}y_B \frac{p_y(y_A, y_B)}{(w_A + y_A)(w_B + y_B)} \\ &= \rho^2 \int_0^\infty \mathrm{d}y_A \int_0^\infty \mathrm{d}y_B \frac{p_y(y_A, y_B)}{(w_A + \bar{y}_A)(w_B + \bar{y}_B)} \\ &\quad \times \left[\sum_{i=0}^\infty \left(\frac{\bar{y}_A - y_A}{w_A + \bar{y}_A} \right)^i \right] \left[\sum_{j=0}^\infty \left(\frac{\bar{y}_B - y_B}{w_B + \bar{y}_B} \right)^j \right] \\ &= \rho^2 \sum_{i,j=0}^\infty (-1)^{i+j} \frac{M_{ij}}{(\bar{y}_A + w_A)^{i+1} (\bar{y}_B + w_B)^{j+1}} , \end{aligned} \tag{B.1}$$

where M_{ij} are the "centered" moments of p_y:

$$M_{ij} \equiv \int_0^\infty \mathrm{d}y_A \int_0^\infty \mathrm{d}y_B p_y(y_A, y_B) \, (y_A - \bar{y}_A)^i (y_B - \bar{y}_B)^j . \tag{B.2}$$

The centered moments can be expressed in terms of the "un-centered" ones, defined as

$$\mathcal{M}_{ij} \equiv \int_0^\infty \mathrm{d}y_A \int_0^\infty \mathrm{d}y_B p_y(y_A, y_B) \, y_A^i y_B^j = (-1)^{i+j} Y^{(i,j)}(0,0) . \tag{B.3}$$

Here $Y^{(i,j)}(0,0)$ is the i-th partial derivative on s_A and j-th partial derivative on s_B of $Y(s_A, s_B)$, evaluated at $(0,0)$. These, in turn, can be expressed as derivatives of Q. For the first terms we have

$$Y^{(0,0)}(0,0) = Y(0,0) = 1 , \tag{B.4}$$
$$Y^{(1,0)}(0,0) = \rho Q^{(1,0)}(0,0) , \tag{B.5}$$
$$Y^{(0,1)}(0,0) = \rho Q^{(0,1)}(0,0) , \tag{B.6}$$
$$Y^{(2,0)}(0,0) = \rho Q^{(2,0)}(0,0) + \rho^2 \left[Q^{(1,0)}(0,0) \right]^2 , \tag{B.7}$$
$$Y^{(1,1)}(0,0) = \rho Q^{(1,1)}(0,0) + \rho^2 \left[Q^{(1,0)}(0,0) \right] \left[Q^{(0,1)}(0,0) \right] , \tag{B.8}$$
$$Y^{(0,2)}(0,0) = \rho Q^{(0,2)}(0,0) + \rho^2 \left[Q^{(0,1)}(0,0) \right]^2 . \tag{B.9}$$

Finally, the derivatives of Q can be evaluated as

$$Q^{(i,j)}(0,0) = (-1)^{i+j} S_{ij} , \tag{B.10}$$

where S_{ij}, we recall, is given by Eq. (57). Note that $S_{01} = S_{10} = 1$ because of the normalization of w_A and w_B, and thus, as already anticipated, $\bar{y}_A = \bar{y}_B = \rho$. In summary, we find

$$M_{00} = 1 , \tag{B.11}$$
$$M_{10} = M_{01} = 0 , \tag{B.12}$$
$$M_{20} = \mathcal{M}_{20} - (\mathcal{M}_{10})^2 = \rho S_{20} , \tag{B.13}$$
$$M_{11} = \mathcal{M}_{11} - \mathcal{M}_{10} \mathcal{M}_{01} = \rho S_{11} , \tag{B.14}$$
$$M_{02} = \mathcal{M}_{20} - (\mathcal{M}_{01})^2 = \rho S_{20} . \tag{B.15}$$

We stress that, in general, it is not true that $M_{ij} = \rho S_{ij}$ (more complex expressions are encountered for higher order terms; cf. the last term in Eq. (56)). Finally, we can write the expansion of $C(w_A, w_B)$:

$$\begin{aligned} C(w_A, w_B) \simeq{} & \frac{\rho^2}{(\rho + w_A)(\rho + w_B)} + \frac{\rho^3 S_{20}}{(\rho + w_A)^3 (\rho + w_B)} \\ & + \frac{\rho^3 S_{11}}{(\rho + w_A)^2 (\rho + w_B)^2} + \frac{\rho^3 S_{02}}{(\rho + w_A)(\rho + w_B)^3} . \end{aligned} \tag{B.16}$$

This is precisely Eq. (58). Using the same technique and a little more perseverance, we can also obtain higher order terms. In particular, Table B.1 reports the moments M_{ij} defined in Eq. (B.2) up to the forth order. This table, together with Eq. (B.1), can be used to write an accurate moment expansion of $C(w_a, w_B)$.

Appendix C: Properties of the Laplace transform

For the convenience of the reader, we summarize in this appendix some useful properties of the Laplace transform. Proofs of the results stated here can be found in any advanced analysis book (**?**, e.g.)]Arfken. Although in this paper we have been dealing mainly with Laplace transforms of two-argument functions, we write the properties below for the case of a function of a single argument for two main reasons: (i) The generalization to functions of several arguments is in most cases trivial; (ii) Several properties can be better understood in the simpler case considered here.

Suppose that a function $f(x)$ of a real argument x is given. Its Laplace transform is defined as

$$\begin{aligned} \mathcal{L}[f](s) &\equiv \lim_{x \to 0^-} \int_x^\infty \mathrm{d}x' \, f(x') \mathrm{e}^{-sx'} \\ &= \int_{0^-}^\infty \mathrm{d}x \, f(x) \mathrm{e}^{-sx} . \end{aligned} \tag{C.1}$$

Note that we use 0^- as lower integration limit in this definition.

The Laplace transform is a *linear operator*; hence, if α and β are two real numbers and $g(x)$ is a function of real argument x, we have $\mathcal{L}[\alpha f + \beta g] = \alpha \mathcal{L}[f] + \beta \mathcal{L}[g]$.

The Laplace transform of the derivative of f can be expressed in terms of the Laplace transform of f. In particular, we have

$$\mathcal{L}\bigl[f'\bigr](s) = s\mathcal{L}[f](s) - f(0^-) \ . \tag{C.2}$$

This equation can be generalized to higher order derivatives. Calling $f^{(n)}$ the n-th derivative of f, we have

$$\mathcal{L}\bigl[f^{(n)}\bigr](s) = s^n \mathcal{L}[f](s) - \sum_{i=0}^{n-1} s^{n-i-1} f^{(n)}(0^-) \ . \tag{C.3}$$

Surprisingly, this equation holds if, for n negative, we consider $f^{(n)}$ to be the $-n$-th integral of f; note that in this case the summation disappears. Hence, for example, we have

$$\mathcal{L}\biggl[\int_{0^-}^{x} f(x') \, \mathrm{d}x'\biggr](s) = \frac{1}{s}\mathcal{L}[f](s) \ . \tag{C.4}$$

Often, properties of the Laplace transform comes in pairs: For every property there is a similar one where the role of f and $\mathcal{L}[f]$ are swapped. Here is the “dual” of property (C.2):

$$\mathcal{L}\bigl[x f(x)\bigr](s) = -\frac{\mathrm{d}\mathcal{L}[f](s)}{\mathrm{d}s} \ , \tag{C.5}$$

or, more generally,

$$\mathcal{L}\bigl[x^n f(x)\bigr](s) = (-1)^n \frac{\mathrm{d}^n \mathcal{L}[f](s)}{\mathrm{d}s^n} \ . \tag{C.6}$$

A similar equation holds for “negative” derivatives, i.e. integrals of the Laplace transform. In this case, however, it is convenient to change the integration limits to (s, ∞). In summary, we can write

$$\mathcal{L}\bigl[f(x)/x\bigr](s) = \int_{s}^{\infty} \mathcal{L}[f](s') \, \mathrm{d}s' \ . \tag{C.7}$$

Given a positive number a, the Laplace transform of the function f shifted by a is given by

$$\mathcal{L}\bigl[f(x-a)\mathrm{H}(x)\bigr](s) = \mathcal{L}[f](s)\mathrm{e}^{-sa} \ , \tag{C.8}$$

where H is the Heaviside function defined in Eq. (30). A dual of this property can also be written:

$$\mathcal{L}\bigl[f(x)\mathrm{e}^{bx}\bigr](s) = \mathcal{L}[f](s-b) \ . \tag{C.9}$$

Finally, we consider two useful relationships between limiting values of f and $\mathcal{L}[f]$:

$$\lim_{x \to 0^+} f(x) = \lim_{s \to \infty} s\mathcal{L}[f](s) \ , \tag{C.10}$$

$$\lim_{x \to \infty} f(x) = \lim_{s \to 0^+} s\mathcal{L}[f](s) \ . \tag{C.11}$$